ГЕОДЕЗИЯ

УДК 528.063

К ВОПРОСУ ОБОСНОВАНИЯ ОБОБЩЁННОГО МЕТОДА НАИМЕНЬШИХ КВАДРАТОВ

д-р техн. наук, проф. В.И. МИЦКЕВИЧ, А.Ю. БУДО
(Полоцкий государственный университет)

Историческое развитие математической обработки геодезических измерений связано с именами таких великих учёных, как К. Гаусс, Ф. Гельмерт, Л. Крюгер, П. Лапласс, А. Лежандр. Научные традиции Гаусса в России продолжили А.П. Болотов, А.Н. Савич, И.И. Померанцев, В.В. Витковский и др. Значительный вклад в развитие теоретического фундамента внесли такие видные учёные-геодезисты Беларуси, как И.И. Купчинов и В.В. Попов. Современное развитие науки и техники обеспечивает новые возможности решения задач геодезии и создаёт предпосылки к разработке новых методов их решения, применительно к современным измерительным технологиям и средствам вычислений. Это позволяет выполнять математическую обработку измерений, реализующую не только классические методы, но и другие, основанные на теории исследования операций. Такой подход расширяет возможности практического использования вероятностно-статистических методов математической обработки наблюдений, в наибольшей степени учитывающих разнообразие информации о характере формирования погрешностей измерений. В связи с этим возрастает активность разработки путей реализации теоретических основ путём разработки алгоритмов и программ, обеспечивающих эффективное применение компьютерных технологий обработки различных по виду и точности геодезических измерений. Целью статьи является изучение вопросов по обоснованию обобщённого метода наименьших квадратов, которое дано в наилучшем изложении Ю.В. Кемницем. Практическое использование теорем по обобщённому методу наименьших квадратов связано с обработкой GPS-спутниковых измерений.

Введение. В 1819 – 1826 годах Гаусс создал знаменитый «Большой мемуар», под названием «Теория такой комбинации наблюдений, при которой вычисленные значения искомых неизвестных были бы возможно менее искажены ошибками измерений». В этом выдающемся труде, продолжая дальнейшую разработку своего метода уравнивания и оценку точности результатов измерений, Гаусс дал новое обоснование метода наименьших квадратов (МНК), поставив требование такого сочетания измеренных данных, чтобы мера точности результата была наибольшей.

В методе МНК предполагается, что измеренные величины являются независимыми и такой подход к уравниванию называется классическим МНК. В середине XX века стали появляться исследования по обобщённому методу наименьших квадратов. В 1946 году выдающийся советский математик, академик А.Н. Колмогоров опубликовал работу, в которой дал новое изложение метода наименьших квадратов, применяя общие понятия линейной алгебры и её геометрической интерпретации.

В 1967 году академик Ю.В. Линник опубликовал свой фундаментальный труд «Метод наименьших квадратов и основы теории обработки наблюдений», в котором широко использовал математический аппарат линейной алгебры. Именно линейная алгебра легла в основу обоснования обобщённого метода наименьших квадратов для случая обработки зависимых результатов измерений.

Нами выполнено обоснование обобщённого метода наименьших квадратов по методике, предложенной Ю.В. Кемницем.

1. Основная теорема обобщенного метода наименьших квадратов

Если 1) N-мерный вектор $T^{изм}$ результатов измерений, вводимых в уравнивание параметрическим методом, имеет как угодно зависимые составляющие, свободные от систематических ошибок, т.е.

$$\mathrm{M}\left(T^{изм}\right) = T^{ур}, \tag{1}$$

где $T^{ур}$ – постоянный N-мерный вектор неизвестных измеренных величин;

2) известны матрица R_1 коэффициентов корреляции этого вектора и его диагональная матрица весов (p);

3) заданы начальные уравнения:

$$AX + T^{выч} = T^{ур}, \tag{2}$$

где A – постоянная матрица коэффициентов параметрических уравнений размера $N \times t$ и ранга $t < N$; X – постоянный t-мерный вектор косвенно отыскиваемых величин; $T^{выч}$ – постоянный N-мерный вектор исходных данных, то оценка x вектора X, удовлетворяющая принципу наибольшего веса, может быть найдена из уравнения [1]:

$$A^T p^{\frac{1}{2}} R_l^{-1} p^{\frac{1}{2}} Ax + A^T p^{\frac{1}{2}} R_l^{-1} p^{\frac{1}{2}} L = 0, \tag{3}$$

где

$$L = T^{выч} - T^{изм}. \tag{4}$$

Корреляционная матрица оценки x имеет вид:

$$K_x = \mu^2 \left(A^T p^{\frac{1}{2}} R_L^{-1} p^{\frac{1}{2}} A \right)^{-1}, \tag{5}$$

где

$$\mu^2 = \frac{\upsilon^T p^{\frac{1}{2}} R_l^{-1} p^{\frac{1}{2}} \upsilon}{n-1} \tag{6}$$

– оценка дисперсии σ^2 единицы веса, причем

$$\upsilon = Ax + L \tag{7}$$

– вектор поправок к составляющим вектора $T^{изм}$.

Доказательство. Учитывая, что матрица R_L симметрична, разложим ее на произведение двух треугольных матриц, транспонированных относительно друг друга:

$$R_L = \Omega^T \Omega, \tag{8}$$

где Ω – правотреугольная матрица.

Введем неособенное линейное преобразование:

$$l = GL, \tag{9}$$

матрица преобразования которого имеет вид:

$$G = \left(\Omega^{-1}\right)^T p^{\frac{1}{2}}. \tag{10}$$

Тогда корреляционная матрица K_l вектора l выразится через корреляционную матрицу K_L вектора L равенством:

$$K_l = G K_L G^T. \tag{11}$$

Поскольку корреляционная матрица K_L связана с матрицей коэффициентов корреляции R_L равенством

$$K_L = \sigma^2 p^{-\frac{1}{2}} R_L p^{-\frac{1}{2}}, \tag{12}$$

где R_L – корреляционная матрица, содержащая на диагонали единицы, а недиагональные элементы равны коэффициентам корреляции.

Подставляя значение K_L в (11) и учитывая разложение (8), приходим к выражению:

$$K_l = \sigma^2 E, \tag{13}$$

из которого заключаем, что линейное преобразование (9) переводит вектор L в вектор l с некоррелированными и равноточными составляющими.

Теперь подвергнем тому же линейному преобразованию с матрицей G матрицу A и векторы $T^{выч}$, $T^{ур}$:

$$A_1 = GA;\ T_1^{выч} = GT^{выч};\ T_1^{ур} = GT^{ур}. \tag{14}$$

Тогда, умножив слева обе части равенства (2) на матрицу G, придем к новым начальным уравнениям:

$$A_1 X + T_1^{выч} = T_1^{ур}, \tag{15}$$

в которых вектору $T_1^{ур}$ можно сопоставить вектор *l*, определяемый равенством (9), как вектор равноточных и некоррелированных результатов измерений.

В таком случае, приняв $P = E$, согласно начальным уравнениям (15), можем написать нормальные уравнения:

$$A_1^T A_1 x + A_1^T l = 0\,, \tag{16}$$

где x – оценка вектора X, удовлетворяющая принципу наибольшего веса.

Подставив в (16) матрицу A_1 из первого равенства (14), вектор l из (9), а после этого матрицу преобразования G из (10) и учтя разложение (8), придем к нормальным уравнениям (3), что и доказывает первую часть теоремы.

Для доказательства формулы (5), приняв $P = E$, формуле для μ придадим вид:

$$K_x = \sigma^2 \left(A_1^T A_1\right)^{-1}.$$

Подставив сюда значение матрицы A_1 из (14) и (8), придем к формуле (5), заменив σ^2 на оценку μ^2.

Проведем доказательство выражения (6). Для этого, приняв $p = E$, по формуле для μ придадим вид:

$$\mu^2 = \frac{\upsilon^T \upsilon}{n-1}\,. \tag{17}$$

Ясно, что вектору l соответствует вектор поправок

$$\upsilon_1 = G\upsilon\,, \tag{18}$$

подставляя который в (17) и учитывая (10) и (8), приведем к формуле (6), что и доказывает теорему полностью.

Можно показать, что если, кроме того, задана некоторая функция

$$\Phi_X = f_0 + f^T X\,,$$

где число f_0 – постоянная, а f – постоянный t-мерный вектор, то удовлетворяющая принципу наибольшего веса ее оценка имеет вид:

$$\Phi_x = f_0 + f^T x\,, \tag{19}$$

где вектор x находят из уравнения (3) [1].

Оценку дисперсии оценки (19) производят при этом по формуле:

$$m_\Phi^2 = \mu^2 f^T \left(A^T p^{\frac{1}{2}} R_l^{-1} p^{\frac{1}{2}} A\right)^{-1} f\,,$$

где μ^2 по-прежнему определяют из выражения (6).

Для оценки надежности эмпирической средней квадратической ошибки единицы веса μ можно, как и ранее, воспользоваться формулой $\sigma_\mu \approx \frac{\mu}{\sqrt{2r}}$ [2].

Обзор формул (3) – (7) обнаруживает, что при вычислениях по ним наибольшая затрата труда падает на обращение матрицы коэффициентов корреляции R_L. Чтобы по возможности избежать этого затруднения, надо вести уравнивание не по параметрическому методу, а по коррелатному, где необходимость такого обращения отпадает [1].

2. Обобщённый коррелатный способ уравнивания

Рассмотрим этот вопрос с позиции принципа наибольшего веса.

Из теории метода наименьших квадратов с позиции принципа наибольшего веса известно, что начальные уравнения (2) можно преобразовывать к условным уравнениям:

$$BT^{ур} + b = 0\,, \tag{20}$$

число которых будет $r = N - t$, причем здесь B – постоянная матрица размера $r \times N$ и ранга $r < N$; b – вспомогательный вектор.

Если вместо $T^{ур}$ в условные уравнения (20) подставить вектор результатов измерений $T^{изм}$, то в правой части равенства (20) вместо нулевого вектора появится так называемый *r*-мерный вектор *w* невязок условных уравнений:

$$BT^{изм} + b = w. \tag{21}$$

Чтобы его устранить, вектору результатов измерений $T^{изм}$ придаем вектор поправок *v*:

$$\hat{T} = T^{изм} + v. \tag{22}$$

и поэтому замена $T^{изм}$ на $\hat{T}$ в (21) дает возможность написать:

$$B\hat{T} + b = 0.$$

Учитывая здесь (21) и (22), приходим к условным уравнениям поправок:

$$Bv + w = 0. \tag{23}$$

Если на основании (18) написать

$$v = G^{-1}v_1. \tag{24}$$

и в (23) произвести соответствующую подстановку вектора *v*, то получим новые условные уравнения поправок:

$$U\text{v}_1 + w = 0, \tag{25}$$

$$U^T = \left(G^{-1}\right)^T B. \tag{26}$$

В этих условных уравнениях поправок участвует вектор v_1 поправок к равноточным и некоррелированным результатам измерений (9). Но в этом случае, как известно из теории метода наименьших квадратов, для нахождения поправок, соответствующих принципу наибольшего веса, надо решить нормальные уравнения коррелат:

$$UU^T k + w = 0, \tag{27}$$

где *k* есть *r*-мерный вектор коррелат, и тогда поправки будут получены из уравнения:

$$v_1 = U^T k. \tag{28}$$

Учитывая теперь (8), (10) и (26), уравниванию (27) придаем вид:

$$Bp^{-\frac{1}{2}} R_L p^{-\frac{1}{2}} B^T k + w = 0. \tag{29}$$

Теперь на основании (8), (10), (24), (26) и (28) получаем

$$v = p^{-\frac{1}{2}} R_L p^{-\frac{1}{2}} B^T k. \tag{30}$$

Уравнения (22), (28) и (30) позволяют найти оценку *x* вектора $T^{ур}$, соответствующую принципу наибольшего веса.

Можно показать [2], что оценка

$$\Phi_x = f_0 + f^T x, \tag{31}$$

где f_0 – постоянное число, а *f* – постоянный *n*-мерный вектор, любой постоянной Φ_x, связанной с вектором *X* уравниванием:

$$\Phi_X = f_0 + f^T X, \tag{32}$$

также будет удовлетворять принципу наибольшего веса, т.е. обладать минимальной дисперсией.

На этом заканчивается решение задачи уравнивания по методу условных уравнений.

Для решения задачи апостериорной оценки точности напишем линейное преобразование

$$\xi = Gx, \tag{33}$$

соответствующее (9) и (18), где ξ есть N-мерный вектор уравненных значений равноточных некоррелированных результатов измерений, являющихся составляющими вектора l.

Согласно (33) можем написать:

$$x = G^{-1}\xi$$

и преобразовать оценку (19) к виду:

$$\Phi_x = f_0 + \varphi^T \xi, \tag{34}$$

где

$$\varphi = \left(G^{-1}\right)^T f\ . \tag{35}$$

Оценка (33) соответствует случаю уравнивания равноточных и некоррелированных результатов измерений. Поэтому, как известно из теории метода наименьших квадратов, для ее эмпирической дисперсии, согласно нормальным уравнениям (27), можем написать выражение:

$$m_\Phi^2 = \mu^2 \left(U^T q + \varphi\right)^T \left(U^T q + \varphi\right), \tag{36}$$

где q есть n-мерный вектор переходных множителей, который находится из решения уравнения:

$$UU^T q + U\varphi = 0\ , \tag{37}$$

а оценка эмпирической дисперсии единицы веса μ получается по формуле:

$$\mu^2 = \frac{v_1^T v_1}{r}\ . \tag{38}$$

Учитывая (8), (10), (26) и (35), можем преобразовать равенства (36) – (38) так:

$$m_\Phi^2 = \mu^2 \left(B^T q + f\right)^T p^{-\frac{1}{2}} R_L p^{-\frac{1}{2}} \left(B^T q + f\right); \tag{39}$$

$$Bp^{-\frac{1}{2}} R_L p^{-\frac{1}{2}} B^T q + Bp^{-\frac{1}{2}} R_L p^{-\frac{1}{2}} f = 0; \tag{40}$$

$$\mu^2 = \frac{v_1^T p^{\frac{1}{2}} R_L^{-1} p^{\frac{1}{2}} v}{r}\ . \tag{41}$$

Ясно, что формулой (41) можно воспользоваться лишь в случае, когда матрица R_L обращена. Чтобы избежать этого неудобства, преобразуем формулу.

Умножив обе части равенства (28) слева на v_1^T и учтя (25), придем к равенству:

$$v_1^T v_1 = -w^T k\ ,$$

с помощью которого формуле (38) придаем вид:

$$\mu^2 = -\frac{w^T k}{r}\ , \tag{42}$$

который и будет желаемым представлением формулы (41), так как вектор невязок легко вычисляется по формуле (21).

В результате всего сказанного можем сформулировать следующую теорему.

3. Вспомогательная теорема обобщенного метода наименьших квадратов

Если: 1) N-мерный вектор $T^{изм}$ результатов измерений, вводимых в уравнивание коррелатным методом условных измерений, имеет как угодно зависимые составляющие, свободные от систематических ошибок, т.е. соблюдается условие (1);

2) известна матрица R_L коэффициентов корреляции этого вектора и его диагональная матрица весов (p);

3) заданы условные уравнения (20) и функция (32), то удовлетворяющая принципу наибольшего веса оценка последней будет (31), где вектор x находится последовательным применением формул (22), (29), (30).

При этом несмещенная оценка дисперсии оценки (31) может быть найдена последовательным применением формул (42), (40) и (39).

В этом методе уравнивания может быть предложен следующий порядок вычислений, который изложим применительно к уравниванию опорных геодезических сетей [1]:

1) на основании схемы опорной сети, на которой указаны, какие элементы ее измерены, составляются условные уравнения (20);

2) подсчитываются невязки условных уравнений (21);

3) по заданным матрицам p и R_L вычисляется вспомогательная матрица:

$$R_p = p^{-\frac{1}{2}} R_L p^{-\frac{1}{2}};$$

4) составляются и решаются нормальные уравнения коррелат (29), которые с помощью матрицы R_p представляют в виде $BR_pB^Tk + w = 0$. При этом как промежуточный результат отдельно фиксируется матрица R_pB^T;

5) вычисляются в соответствии с (30) поправки и уравненные значения (22) результатов измерений;

6) по уравненным значениям результатов измерений обычным порядком вычисляются необходимые элементы уравниваемой опорной сети: дирекционные углы сторон, координаты пунктов т.п. На этом решение задачи уравнивания заканчивается;

7) для решения задачи апостериорной оценки точности одновременно с составлением условных уравнений выбираются элементы сети, подлежащие оценке точности, и для них составляются функции типа (32). Если выбранные элементы сети выражаются через составляющие вектора $T^{ур}$ нелинейно, то предварительно производится известным образом линеаризация соответствующих функций;

8) попутно с составлением нормальных уравнений коррелат составляются и решаются уравнения (40):

$$BR_pB^Tq + BR_pf = 0 .$$

Число таких систем уравнений равно числу элементов сети, подлежащих оценке точности;

9) по формуле

$$\mu = \sqrt{-\frac{w^Tk}{r}}$$

вычисляется эмпирическая средняя квадратическая оценка единицы веса. Оценка ее надежности может быть осуществлена по приближенной формуле:

$$m_\mu = \frac{m}{\sqrt{2r}};$$

10) наконец, по формуле

$$m_F = \mu\sqrt{\left(B^Tq + f\right)^T R_p \left(B^Tq + f\right)}$$

производится оценка точности уравновешенных значений выбранных элементов сети.

Числовой пример на уравновешивание зависимых результатов измерений методом условных измерений можно найти в [3].

Основная и вспомогательная теоремы обобщенного метода наименьших квадратов касаются лишь уравнивания классическими методами. Существенного труда не составляет получить соответствующие теоретические факты и для других методов уравнивания (например, параметрический способ с дополнительными условиями, связанный условиями, и т.д.). Однако принципиально новых научных положений при этом получено не будет, так как все эти методы уравнивания путем чисто алгебраических преобразований могут быть сведены к двум методам, рассмотренным в этом исследовании.

Таково обоснование обобщенного метода наименьших квадратов на принципе наибольшего веса. Можно получить обоснование обобщенного метода наименьших квадратов на принципе максимального правдоподобия, допустив результаты измерений нормально распределенными. Для этого может быть использован прием, предложенный в [4].

Заключение

Исследования показали, что изложенные выше теоремы позволяют дать теоретический фундамент к обобщённому методу *Lp*-оценок, в котором используются степени n (при n = 2,0 будет обобщённый МНК; при n = 1,0 – обобщённый метод наименьших модулей, который ещё не рассматривался в научной литературе).

Практические приложения обобщённого метода *Lp*-оценок не могут быть применены без доказательств соответствующих теорем. Эти теоремы будут логическим продолжением сделанных в настоящей статье доказательств по их практическому применению. Последнее наиболее эффективно при обработке зависимых результатов измерений в относительном методе спутниковых *GPS* / ГЛОНАСС технологий.

ЛИТЕРАТУРА

1. Кемниц, Ю.В. Математическая обработка зависимых результатов измерений / Ю.В. Кемниц. – М.: Недра, 1970. – 192 с.
2. Кемниц, Ю.В. Обобщённая формула средней квадратической ошибки нелинейной функции / Ю.В. Кемниц // Изв. вузов. Геодезия и аэрофотосъёмка. – 1958. – № 5.
3. Гордеев, Ю.А. О применении принципа наименьших квадратов при уравновешивании зависимых результатов измерений / Ю.А. Гордеев // Изв. вузов. Геодезия и аэрофотосъёмка. – 1960. – № 2.
4. Линник, Ю.В. Метод наименьших квадратов и основа математической теории обработки наблюдений / Ю.В. Линник. – М.: Физматгиз, 1962. – 349 с.

Поступила 22.03.2007